Exercise 1: Gaussian Processes - Prediction

Let $\mathcal{X} = \mathbb{R}$ and assume the following statistical model

$$y = f(x) + \epsilon, \qquad \epsilon \sim \mathcal{N}\left(0, \sigma^2\right),$$

where $f(x) \in \mathcal{GP}\left(0, k(x, x')\right)$. Suppose the covariance function of the GP is

$$k(x, x') = \mathbb{1}_{[|x-x'|<1]} \cdot (1 - |x - x'|)$$

and we have seen the training data:

i	$\mathbf{x}^{(i)}$	$y^{(i)}$
1	1.6	3.0
2	2.8	3.3
3	0.5	2.0
4	3.9	2.7

As a test input we observe $x_* = 1.2$. Recall that the predictive distribution for $f(x_*)$ is

$$f(x_*) \mid \mathbf{X}, \boldsymbol{y}, x_* \sim \mathcal{N}(m_{\text{post}}, k_{\text{post}}).$$

with

$$\begin{aligned} m_{\text{post}} &= \boldsymbol{K}_*^T \left(\mathbf{K} + \sigma^2 \cdot \boldsymbol{I}\right)^{-1} \boldsymbol{y} \\ k_{\text{post}} &= K_{**} - \boldsymbol{K}_*^T \left(\mathbf{K} + \sigma^2 \cdot \boldsymbol{I}\right)^{-1} \boldsymbol{K}_*, \end{aligned}$$

Here, $\mathbf{K} = \left(k\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right)\right)_{i,j}$, $\boldsymbol{K}_* = \left(k\left(x_*, \mathbf{x}^{(1)}\right), ..., k\left(x_*, \mathbf{x}^{(n)}\right)\right)^\top$ and $K_{**} = k(x_*, x_*)$.

(a) Compute the predictive mean m_{post}.

(b) Compute the predictive variance k_{post}.

(c) Repeat the calculations from (a) and (b) by using as the test input $x_* = \mathbf{x}^{(i)}$ for each $i = 1, 2, 3, 4$, respectively.

(d) Based on your calculations so far, try to sketch the posterior Gaussian process.

(e) If the nugget σ^2 would be zero, how would the posterior Gaussian process (roughly) look like?